Chapter 26

ARResT/Interrogate Immunoprofiling Platform: Concepts, Workflows, and Insights

Nikos Darzentas

Abstract

ARResT/Interrogate was built within the EuroClonality-NGS working group to meet the challenge of developing and applying assays for the high-throughput sequence-based profiling of immunoglobulin (IG) and T-cell receptor (TR) repertoires. We herein present basic concepts, outline the main workflow, delve into EuroClonality-NGS-specific aspects, and share insights from our experiences with the platform.

Key words Immunoglobulin, Antigen receptors, IG, TR, Bioinformatics, Pipeline, Sequence analysis

1 Introduction

Immunoglobulins (IG) and T-cell receptors (TR) are highly adaptive molecular receptors involved in antigen recognition and enormously variable immunological responses. The advent of sequence-based profiling of IG and TR repertoires has been instrumental for understanding such responses, both normal and pathologic, the latter encompassing a wide range of diseases with an underlying immune cause. This unprecedented capability has also brought along novel and unique challenges [1]; this chapter will cover the bioinformatic one, from the perspective of the ARResT/Interrogate immunoprofiling platform.

ARResT (abbreviation of Antigen Receptors Research Tool, http://bat.infspire.org/arrest) comprises a handful of tools developed over the years within focused groups. It originated in the days of Sanger sequence analysis toward delineating subsets of stereotyped antigen receptor sequences in chronic lymphocytic leukemia (CLL) [2, 3].

ARResT/Interrogate [http://arrest.tools/interrogate] was built from the grounds up within the EuroClonality-NGS working group [http://euroclonality.org] to initially support the development of the group's NGS assays and eventually to apply them in

research and clinical applications [4, 5]. ARResT/Interrogate is able to: automatically paired-end-join and concatenate input files; use spreadsheet sample sheets to make data and metadata available to itself and the user; identify, tag, trim, and report on primer sequences (and primer dimers); annotate and identify all rearrangement types (or 'junction classes') of all IG/TR loci; offer powerful interactive tools to the user for mining results; identify, filter, and use the EuroClonality-NGS central in-tube quality/quantification control (cIT-QC, or spike-ins) for abundance normalization; generally support EuroClonality-NGS assays, also with bespoke analytical and visual functionalities; and provide detailed logs and feedback to the user.

ARResT/Interrogate will be continuously updated, and therefore bioinformatic and user interface details included herein may not stay the same over time. We advise readers/users to seek the latest information on the ARResT/Interrogate browser [http://arrest.tools/interrogate] and on the EuroClonality-NGS website [http://euroclonality.org]. For the same reason, we chose not to focus on application-specific methods, also because they are covered in other chapters in this book. Still, the general concepts and workflows included in this chapter should be considered safe, as should the Notes below from our years of experience both developing and using ARResT/Interrogate.

1.1 Design

ARResT/Interrogate consists of the pipeline and the browser (its user interface). The browser features four main "panels" for logically organized and ordered steps (Fig. 1):

1. Access to pipeline ("processing").
2. Access to pipeline results ("file").
3. Analysis of immunogenetic features ("questions").
4. Retrieval and analysis of sequences ("forensics") (*see* **Note 1**).

There is also the "HQ" panel that offers introductory text and specific notes and advice (in separate "tabs"). There are more panels to serve special applications, e.g., clonality assessment, but they are by default hidden and may be accessed by switching "user modes" with the widget on the top left (set at "Interrogate.simple" in Fig. 1).
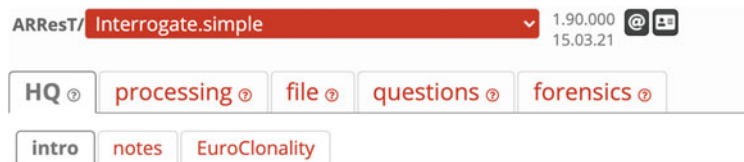

Fig. 1 ARResT/Interrogate browser (user interface), with "panels," "tabs," and the "user mode" selection widget

1.2 Primers

ARResT/Interrogate is able to identify, tag, trim, and report on primer sequences (and primer dimers), including making the results available for a fully interactive analysis. The trimming allows for less artificial sequence data to be processed more accurately and more efficiently, while the reporting allows for the primer-based results to be directly used for quality control and development. *See* also **Notes 2** and **3** about trimming.

1.3 Rearrangements and Junctions

ARResT/Interrogate is able to annotate and identify all rearrangement types of all IG/TR loci. We call these rearrangement types "junction classes." They include "complete," e.g., IG's VJ:Vh-(Dh)-Jh; "incomplete," e.g., TR's DJ:Db-Jb; and "other," e.g. IG's Vk-Kde or intron-Kde (Table 1). For junction classes with no biologically relevant junctional anchors (i.e., residues that define the CDR3 region, as per IMGT), we decided to introduce virtual ones—this enables consistent and informative results across all junction classes, assisting the user to focus on the most variable part of the rearrangement. For the D genes in DJ, VD, and DD incomplete junction classes, we use recombination signal sequence (RSS) heptamers: the last triplet of the heptamer in 5′ and the first triplet of the heptamer in 3′. For the intron RSS in the IGK locus, we use a CCC triplet between the EuroClonality-NGS primer and the RSS heptamer, while for Kde in the IGK locus the final triplet after the RSS heptamer and before the EuroClonality-NGS primer is used. In the majority of cases, these anchors are far enough from the junctional point to allow for nucleotide trimming without affecting their presence, but ARResT/Interrogate is anyway able to report rearrangements even with the anchors trimmed or mutated. This is also true for normal anchors in complete rearrangements.

Anchors overview:

5′ side of junction.

V genes: C aa = TG[CT] nt.

D genes: V aa = GT[any] nt, the last triplet of the 5′ heptamer.

intron: P aa = CCC nt, a triplet between primer and heptamer.

3′ side of junction.

J genes: W aa = TGG nt or F aa = TT[CT] nt.

D genes: H aa = CA[CT] nt, the first triplet of the 3′ heptamer.

Kde: R aa = CGA nt, final triplet after heptamer and before primer.

Table 1
Junction classes supported by ARResT/Interrogate and the EuroClonality-NGS amplicon and capture assays

Junction class	Receptor	Chain	Type	EuroClonality-NGS amplicon primer set	EuroClonality-NGS DNA capture (ECNDC)
VJ:Vh-(Dh)-Jh	IG	H	Complete	IGH-VJ	Yes
DJ:Dh-Jh	IG	H	Incomplete	IGH-DJ	Yes
VJ:Vk-Jk	IG	K	Complete	IGK-VJ-Kde (IGK)	Yes
Vk-Kde	IG	K	Other	IGK-VJ-Kde (IGK)	Yes
intron-Kde	IG	K	Other	intron-Kde (IGK)	Yes
VJ:Vl-Jl	IG	L	Complete		Yes
VJ:Va-Ja	TR	A	Complete	TRD	Yes
VJ:Va-(Dd)-Jd	TR	D	Complete	TRD	Yes
VJ:Vd-(Dd)-Ja	TR	D	Complete	TRD	Yes
DJ:Dd-Ja	TR	D	Incomplete	TRD	Yes
VD:Va-Dd3	TR	D	Incomplete		Yes
VJ:Vb-(Db)-Jb	TR	B	Complete	TRB-VJ	Yes
DJ:Db-Jb	TR	B	Incomplete	TRB-DJ	Yes
VD:Vb-Db	TR	B	Incomplete		Yes
DD:Db-Db	TR	B	Other		Yes
VJ:Vd-(Dd)-Jd	TR	D	Complete	TRD	Yes
DJ:Dd2-Jd	TR	D	Incomplete	TRD	Yes
VD:Vd-Dd3	TR	D	Incomplete	TRD	Yes
DJ:Dd3-Jd	TR	D	Incomplete		Yes
VD:Vd-Dd2	TR	D	Incomplete		Yes
DD:Dd2-Dd3	TR	D	Other	TRD	Yes
DD:Dd3-Dd2	TR	D	Other		Yes
VJ:Vg-Jg	TR	G	Complete	TRG	Yes

2   Materials

2.1   Sequence Input

1. All sequences are uploaded through the "processing" panel.
2. Sample sequences should be uploaded in FASTQ (preferably) or FASTA format, preferably also compressed in "gunzip" format (extension ".gz"). Also *see* **Note 4**.

3. Primer or tracer sequences should be uploaded in uncompressed FASTA format.

4. Filenames should not contain spaces or any special characters; underscores and hyphens are allowed (in fact, encouraged for clarity). This is—generally speaking—a good advice for any files to be used with bioinformatic tools.

5. ARResT/Interrogate can automatically recognize forward/reverse and multilane sequence files, the former being paired-end-joined and the latter concatenated. Since our code is based on Illumina, such files should contain "_L001_R1_" incremented accordingly. The user will be alerted if we believe there are issues with this logic, e.g., if we only have forward (R1) and not reverse (R2).

6. There are more checks on files, including for bad format, zero size, etc.; the user should watch out for relevant pipeline feedback.

2.2 Availability, Requirements, Contact

ARResT/Interrogate is currently available online at arrest.tools/interrogate (*see* **Note 5**); therefore, compute and storage requirements on the user side are limited. We nevertheless urge the use of a modern computer and web browser. In case of trouble using ARResT/Interrogate, please email contact@arrest.tools with as many details as possible, on what was done and if the issue persisted after a fresh start. Screenshots are invaluable, even if the browser has crashed and is grayed out.

2.3 Sample Sheet

1. We will mention sample sheets a number of times below. A sample sheet, a spreadsheet in Microsoft Excel format, is uniquely useful to provide the pipeline and the browser (and the user) with data and metadata that can run different pipeline options for different samples, provide cell counts (deduced from amount of DNA) for spike-in-based normalization, and help select/filter/order/rename/identify user's samples in the browser.

2. The ARResT/Interrogate sample sheet offers a number of predefined columns (i.e., ARResT/Interrogate expects these column names for the information to be used properly) and the possibility to add many others with flexible column names.

3. The most important predefined columns (again, do not change the column names or use them for other purposes) are:

   (a) Sample: required—unique for every sample and part or whole of the sample's sequence filenames.

   (b) Cells: number of cells, based on amount of DNA of, e.g., patient, to be used for quantification.

sample	cells	primer set	primers	rearrangements	scenario	tracers	patient	entity	timepoint
pt1_IGH-VJ-FR1_S1_L001_R1_001	15000	IGH-VJ-FR1	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_IGH-DJ_S2_L001_R1_001	15000	IGH-DJ	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_IGK-VJ-Kde_S3_L001_R1_001	15000	IGK-VJ-Kde	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_intron-Kde_S4_L001_R1_001	15000	intron-Kde	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_TRB-VJ_S5_L001_R1_001	15000	TRB-VJ	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_TRB-DJ_S6_L001_R1_001	15000	TRB-DJ	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_TRD_S7_L001_R1_001	15000	TRD	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt1_TRG_S8_L001_R1_001	15000	TRG	EC-NGS.set	all	EC-NGS.marker_identification		pt1	B-ALL	diagnosis
pt2_IGH-VJ-FR1_S9_L001_R1_001	100000	IGH-VJ-FR1	IGH-leader.set	IGH	routine.follow_up	marker.fasta	pt2	CLL	post-therapy

Fig. 2 Sample sheet example. Red are predefined columns, and blue are flexible columns

(c) Primer set: please use one of IGH-VJ, IGH-DJ, IGK-VJ-Kde, intron-Kde, TRB-VJ, TRB-DJ, TRD, and TRG.

(d) Primers: name(s) of file(s) of primers.

(e) Scenario: if one wants to run different pipeline scenarios for different samples.

(f) Rearrangements: rearrangement type(s) to be identified for each sample.

(g) Tracers: name(s) of file(s) of tracers (i.e., rearrangements of interest, including spike-ins or artifacts).

(h) Select: which samples should be analyzed, also in batches (i.e., could be 'x' or a batch number).

4. Check the example sample sheet in Fig. 2, in which red are predefined columns and blue are flexible columns.

3   Methods

3.1 A Basic Workflow

1. Visit http://arrest.tools/interrogate (*see* **Note 5**) and log in; this requires an account, which can be requested by emailing contact@arrest.tools.

2. Switch to the "processing" panel.

(a) Create a new analysis or select an existing one, otherwise the "default" will be used, which is OK. Also *see* **Note 6**.

(b) Upload sample sequences in compressed FASTQ/A format (*see* Subheading 2).

(c) The default scenario ("ARResT.profile") should work fine in any case. One may select a different user mode or pipeline scenario, especially when deploying EuroClonality-NGS assays (*see* Subheading 3.2).

(d) One may use own primer sequences by uploading them in uncompressed FASTA format and selecting them under "scenario options" (there are instructions on the user interface). In general, please study primers (e.g., *see* **Notes 3**, **7**, and **8** as to why). Also *see* **Note 9**.

(e) Click on the blue "test it" button when ready; if the test goes well (otherwise please follow the advice in the "process output" tab), click on the green "process" button to start the actual run.

(f) There is no need to wait, one may even close the browser; either log in later or, better, make sure to provide an email address to receive email notifications. Also *see* **Note 10**.

(g) If the run was not successful, the email notification's subject will include "(SOME SAMPLES) FAILED", pay attention to the pipeline feedback as to why, or email contact@arrest.tools.

3. Switch to the "file" panel.

(a) Select results in the drop-down widget, select filtering level (*see* **Note 11**), click "load results".

(b) One may browse the run and sample reports, paying attention to quality control (QC) information, alarms (and our hints and tips for possible causes and solutions), basic numbers like percentage of reads with junction (*see* **Note 7**) that are also color-coded to provide visual feedback. Alarms include:

- Low number of reads "5′ primed in R1" or "3′ primed in R2", indicating wrong or missing primers, noisy reads, i.e., compromised primer alignment, etc.

- Low number of reads "3′ primed in R1" or "5′ primed in R2", indicating long or trimmed amplicons (with FR1 or FR2 primers for example) not covered by the sequenced read length, or wrong or missing primers.

- High number of reads "short"—sequence artifacts are generally an explanation, and if primers were used with the pipeline, one may also see an alarm about primer dimers.

4. Switch to the "questions" panel.

(a) The main series of widgets are split into "select" on the left and "filter" on the right (Fig. 3).

(b) Note that if samples are "QC-failed" (*see* **Note 12**), they will not be available here by default; uncheck appropriate widget in "samples options" to include them back in.
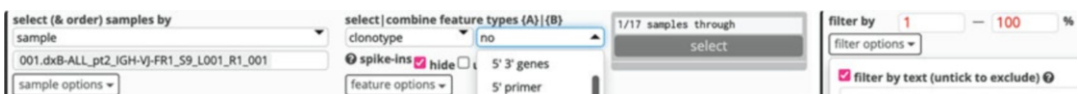
Fig. 3 "Select" and "filter" widgets on the "questions" panel
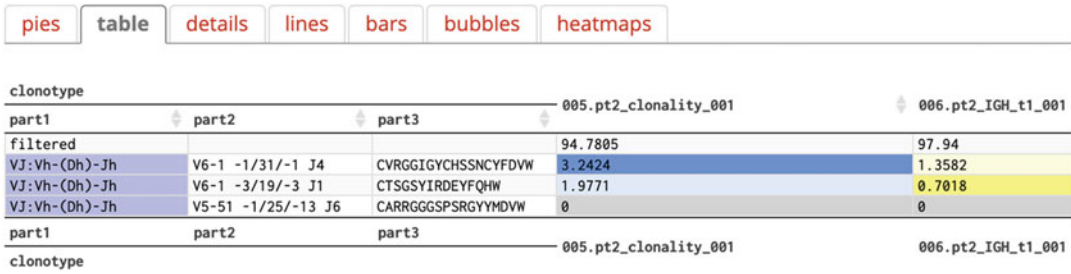
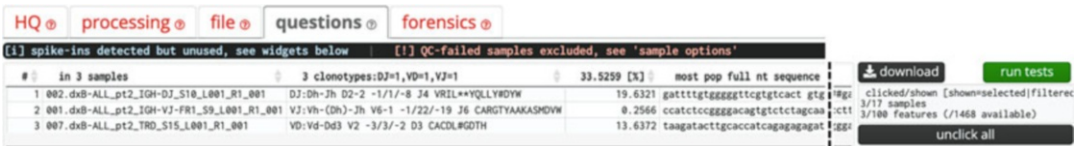**Fig. 4** "Table" visualization of the "questions" panel

Fig. 5 The "minitable" for tabulation and downloading of selected features and their most popular sequences

(c) By default, "clonotypes" will be selected to be shown across samples. One may select one or two feature types among many.

(d) Make sure to click on "select" or "filter" after changing options, in which case these buttons are black (vs. gray).

(e) The panel provides access to multiple tabs for different visualizations, with an interactive table being the default (Fig. 4).

(f) One may click on specific clonotypes, in which case the corresponding information is shown in what is called a "minitable" near the top of the page; importantly, here also the most popular sequence of the clonotypes is provided (Fig. 5). Apart from being able to download this table for reporting, one may also "run tests" on the sequences (see below).

5. Switch to the "forensics" panel, after having clicked on at least one clonotype in "questions."

(a) One may retrieve and download all stored sequences of the clonotype in the "sequences" tab. The sequence variation that will undoubtedly appear in the retrieved sequences could be biological variability (e.g., somatic hypermutation) or technical noise including PCR or sequencing errors or amplification by different primers (*see* **Note 3**). Also, the retrieved sequences are not necessarily all possible sequences from the original sample, as we mainly avoid storing sequences supported by a single read unless they are the only representative of a combination of features.

(b) The "tests" tab (also accessible via "runs tests" in the "questions" panel) offers the possibility to annotate the sequences in different ways. When checking the "Interrogate" option, one will get more color-coded information than what is available in "questions," including D genes and more detailed segmentation. "AssignSubsets" provides access to ARResT/AssignSubsets for assignment of IGH rearrangements to major stereotyped subsets of chronic lymphocytic leukemia (CLL) [http://bat.infspire.org/arrest/assignsubsets] [3].

3.2 EuroClonality-NGS Assays

We will now provide more information on EuroClonality-NGS-specific aspects.

1. EuroClonality-NGS primer sets.

   (a) The EuroClonality-NGS amplicon assay uses eight tubes for the eight EuroClonality-NGS primer sets: IGH-VJ-FR [1–3], IGH-DJ, IGK-VJ-Kde, intron-Kde, TRB-VJ, TRB-DJ, TRD, and TRG.

   (b) It is useful for ARResT/Interrogate to know the primer set of the sample, and therefore, we try to auto-detect it, otherwise the sample is considered "pooled." If the sample is not pooled, these primer set names should be used bookended by _, e.g., sample1_IGH-VJ-FR1_[...] or sample2_IGK-VJ-Kde_[...]. If still detected wrongly, the name should be edited accordingly as to affect the process, either way. Another way is with a sample sheet and its "primer set" column (*see* Subheading 2.3).

   (c) Starting from version 1.90, ARResT/Interrogate specifically tags rearrangements that do not match the primer set (e.g., an VJ:Vg-Jg in an IGK tube) as contamination—one of the advantages of the EuroClonality-NGS assays using one primer set per tube.

2. EuroClonality-NGS central in-tube quality/quantification control (cIT-QC), or spike-ins.

   (a) If one uses spike-ins and wants to access normalized values (i.e., number of cells instead of number of reads), it is also necessary to provide the number of cells (derived from the DNA amount) in the sample, e.g., ~15,000 cells from 100 ng of DNA; this will be used as the denominator for the ratio calculation. There is a widget in the "processing" panel and the same in the "questions" panel (Fig. 6), which sets the value for all samples; if different values need to be set for different samples, this needs to be done with a sample sheet and its "cells" column (*see* Subheading 2.3). Do not include spike-in cells in those numbers.

Fig. 6 Messages and widgets related to cIT-QC (spike-ins)

      (b) One should be able to see extra relevant widgets and messages in "questions" (and remember to hover over the "?" tooltip anchors); to see normalized abundances, check the "use" box (Fig. 6).

3. Pipeline scenarios and browser functionalities for EuroClonality-NGS assays.

      (a) It is important to select the appropriate user modes to properly analyze data from EuroClonality-NGS assays. One of the automations is the preset of appropriate pipeline scenarios in the "processing" panel with the aforementioned primers and spike-ins.

      (b) Switch to the "Interrogate.EC-NGS marker identification" user mode for the assays described in [6]. These assays involve one primer set per tube, plus spike-ins in each tube.

      (c) Switch to the "Interrogate.EC-NGS clonality assessment" user mode for the assays described in [7]. These assays pool the primer set tubes after PCR but before sequencing; therefore, ARResT/Interrogate needs to computationally separate them before calculating abundances. There are currently no spike-ins included. This user mode also enables a bespoke panel, "reporting," in which ARResT/Interrogate separates the different primer sets from the pooled data sample creating one view for each—see the VJ:Vh-(Dh)-Jh and DJ:Dh-Jh views (the latter partially and with a faint red background because of the low number of reads—121, in dark red background—included in it) in Fig. 7.

Fig. 7 Views from the bespoke "reporting" panel of the "Interrogate.EC-NGS clonality assessment" user mode, with two of the pooled primer sets separated, normalized, and presented to the user

4 Notes

1. We strongly advise to spend some time clicking around and hovering over the tooltip markers ("?"), especially after switching to the "Interrogate.advanced" user mode to enable more widgets and options; one should at least be generally aware of what is possible.

2. Although the primer is artificial and may compromise downstream analyses, it sometimes is necessary to keep it on in order to have enough sequence to annotate and thus identify a rearrangement (currently, for EuroClonality-NGS primers, IGH-D, IGK-DE, TRG-J primers are kept on the sequence).

3. Depending on the primer sequences used, one may face potentially confounding situations, especially since ARReST/Interrogate by default trims away the primer sequences.

    Amplification by different primers annealing on the same template may result in slightly different sequences of, e.g., the same clonotype—keep that in mind when looking at combinations of primers and clonotypes, or retrieve sequences of such a clonotype.

    Amplification by different primers annealing on the same template that result in the same sequence and length means that to fully study primers one needs to disable primer trimming so that the sequences remain separate; otherwise, only one primer is remembered per unique sequence that

might not represent the full picture. To do this, enable the "primer_ext" pipeline option, or use the "ARResT.profile.primer_ext" scenario as a template, or email contact@arrest.tools.

4. If the data are too big and/or one is facing upload issues, please email contact@arrest.tools to ask for access to our FTP service (it is planned to make it available by default via the ARResT/Interrogate user interface).

5. If the server (or "station" as we call them) is busy or too slow, please revisit arrest.tools/interrogate to be redirected to a different station—please do not bookmark any final link with specific station numbers.

6. To best analyze and report on markers across diagnostic and follow-up samples (the latter usually coming from a separate, later NGS run), eventually upload all files in the same "analysis" and process together with the same ARResT/Interrogate version.

7. When facing a low percentage of reads with a junction, check the "postmortem" section of the sample report. The first example in Fig. 8 (and below) without a junction has 1181 high-quality and 12 low-quality forward reads and was

Fig. 8 View of the sample report with the "postmortem" section expanded—most abundant examples of sequences with and without junction are shown

bookended by the IGK-INTR-A-1 and IGK-J-A-1 EuroClonality-NGS primers, which actually do not make sense as a pair (IGK intron and IGK J). The second example has reverse reads; it had to go through a more sensitive workflow ("retried") and ended up with an unsafe IGHJ gene assignment ("unsafeJ"), and only had the 5′ IGHV primer on the sequence.

(a) *With no junction, top5 corpses w >=10rds, sorted by weight (fwdHQ:fwdLQ|revHQ:revLQ).*
>1181:12|0:0_____IGK-INTR-A-1_+_IGK-J-A-1__M8KAE:02679:02368.
CACCGCGCTCTTGGGGCAGCCGCCTTGCCGC-TAGTGGCCGTGG[…]
>0:0|136:2____retried;unsafeJ;__IGH-V-FR3-J-1_+_no__M8KAE:01359:02354.
CTCCGTGAAGGGCAGATTCATGAACA-GAATTTTATTGCAGTGTG[…]

8. Primers are useful to safeguard amplicon completeness and thus junction safety. Demand that both primers are present on the sequence of interest if, for example, the lab work is known to produce incomplete amplicons.

9. Do not mix IGH-VJ-FR* primer sets in pipeline options, e.g., FR3 primers will wrongly trim FR1/FR2 reads.

10. Running times may vary heavily, depending on sample number, depth and clonality, read length, and sequence noise.

11. Regarding "results filtering" in the "file" panel, keep in mind to switch to "not pre-filtered" when looking for very low abundance clonotypes.

12. As a "fail" "QC status" in the run report does not necessarily mean that the sample is unusable, one may reinsert it back into the analysis in the "questions" panel. Final decision is with the user, based on context (kind of sample, DNA quality, and purpose); QC status is just meant to attract the user's attention to potential issues.

References

1. Langerak AW, Brüggemann M, Davi F, Darzentas N, van Dongen JJM, Gonzalez D et al (2017) High-throughput Immunogenetics for clinical and research applications in immunohematology: potential and challenges. J Immunol 198:3765–3774

2. Darzentas N, Hadzidimitriou A, Murray F, Hatzi K, Josefsson P, Laoutaris N et al (2010) A different ontogenesis for chronic lymphocytic leukemia cases carrying stereotyped antigen receptors: molecular and computational evidence. Leukemia 24:125–132

3. Bystry V, Agathangelidis A, Bikos V, Sutton LA, Baliakas P, Hadzidimitriou A et al (2015) ARResT/AssignSubsets: a novel application for robust subclassification of chronic lymphocytic leukemia based on B cell receptor IG stereotypy. Bioinformatics 31:3844–3846

4. Bystry V, Reigl T, Krejci A, Demko M, Hanakova B, Grioni A et al (2017) ARResT/

interrogate: an interactive immunoprofiler for IG/TR NGS data. Bioinformatics 33:435–437

5. Knecht H, Reigl T, Kotrová M, Appelt F, Stewart P, Bystry V et al (2019) Quality control and quantification in IG/TR next-generation sequencing marker identification: protocols and bioinformatic functionalities by EuroClonality-NGS. Leukemia 33:2254–2265

6. Brüggemann M, Kotrová M, Knecht H, Bartram J, Boudjogrha M, Bystry V et al (2019) Standardized next-generation sequencing of immunoglobulin and T-cell receptor gene recombinations for MRD marker identification in acute lymphoblastic leukaemia; a EuroClonality-NGS validation study. Leukemia 33:2241–2253

7. Scheijen B, RWJ M, Rijntjes J, van der Klift MY, Möbs M, Steinhilber J et al (2019) Next-generation sequencing of immunoglobulin gene rearrangements for clonality assessment: a technical feasibility study by EuroClonality-NGS. Leukemia 33:2227–2240